The Intersession Process in Psychotherapy for Anorexia Nervosa: Characteristics and Relation to Outcome

Armin Hartmann,[1] Almut Zeeck,[1] Wolfgang Herzog,[2] Beate Wild,[2] Martina de Zwaan,[3] Stephan Herpertz,[4] Markus Burgmer,[5] Joern von Wietersheim,[6] Sefik Tagay,[7] Andreas Dinkel,[8] Bernd Löwe,[9] Gaby Resmark,[10] David Orlinsky,[11] and Stephan Zipfel[10]

[1] *Department of Psychosomatic Medicine and Psychotherapy, Medical Center, University Freiburg, Germany*
[2] *Center for Psychosocial Medicine, Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Germany*
[3] *Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Hannover, Germany*
[4] *Clinic for Psychosomatic Medicine and Psychotherapy, LWL University Hospital of the Ruhr, University of Bochum, Germany*
[5] *Clinic for Psychosomatic Medicine and Psychotherapy, University Hospital Münster, Germany*
[6] *Department of Psychosomatic Medicine and Psychotherapy, University Hospital of Ulm, Germany*
[7] *Clinic for Psychosomatic Medicine and Psychotherapy, LVR Hospital Essen, University of Duisburg-Essen, Germany*
[8] *Clinic for Psychosomatic Medicine and Psychotherapy, University of Technology Munich, Germany*
[9] *Institute and Outpatient Clinic for Psychosomatic Medicine and Psychotherapy, University Hospital Hamburg-Eppendorf, Germany*
[10] *Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tübingen, Germany*
[11] *Department of Comparative Human Development, University of Chicago, IL, USA*

Objective: The "inter session process" (ISP) is defined as therapy-related conscious thoughts, memories, and emotions that patients and therapists experience between psychotherapy sessions. It indicates how the participants process and use treatment. The main aim of this study is to describe the ISP characteristics of patients in outpatient treatment for anorexia nervosa (AN). An additional aim is to explore the relation between patients' ISP and treatment outcome. **Method:** Patients taking part in a randomized controlled trial on outpatient psychotherapy for AN (ANTOP) filled in the Intersession Experience Questionnaire before each of the 40 psychotherapy sessions. Trajectories of different aspects of the ISP were analyzed with growth curve models based on orthogonal polynomials and tested for differences between 3 outcome categories (recovery, partially recovered, full syndrome AN). **Results:** Data from 108 cases were available for analysis. ISP facets showed diverse, mostly nonlinear, trajectories over the course of treatment. Less favorable outcomes were associated with higher levels of patients' experiencing negative emotions when recalling therapeutic dialogue, thinking about therapy during dreaming/drowsy states, and applying therapeutic learning (in the second half of treatment). **Conclusions:** Findings confirm an overall relation between ISP and treatment outcome. In outpatient treatment in AN, patients with a less successful course seem to be more preoccupied with

Armin Hartmann and Almut Zeeck contributed equally to this article.

Please address correspondence to: Armin Hartmann, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Freiburg, Hauptstrasse 8, 79104 Freiburg. E-mail: armin.hartmann@uniklinik-freiburg.de
Almut Zeeck, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Freiburg, Hauptstrasse 8, 79104 Freiburg. E-mail: almut.zeeck@uniklinik-freiburg.de

JOURNAL OF CLINICAL PSYCHOLOGY, Vol. 0(0), 1–19 (2016)
© 2016 Wiley Periodicals, Inc.
Published online in Wiley Online Library (wileyonlinelibrary.com/journal/jclp). DOI: 10.1002/jclp.22293

therapy and therapist between sessions. For the ISP facet "applying therapeutic learning," findings point to an optimal range dependent on treatment phase. Growth curve modeling is required to describe the nonlinear trajectories of ISP facets. © 2016 Wiley Periodicals, Inc. J. Clin. Psychol. 0:1–19, 2016.

Keywords: anorexia nervosa; intersession process; outcome

Introduction

Bohart and Wade (2013) emphasized that "client factors" generally account for the largest amount in variance when explaining outcome. Client or patient factors include sociodemographic features or aspects of psychopathology (Herzog, Schellberg, & Deter, 1997) and the active contribution of patients to the psychotherapeutic process. This active contribution can be conceptualized as (a) taking place *within* a psychotherapy session ("in-session process": e.g., participation, contribution to the therapeutic alliance) and (b) happening between sessions ("intersession process" [ISP]), as conceptualized by Orlinsky and Howard in their generic model of psychotherapy (Orlinsky & Geller, 1993; Orlinsky, Geller, Tarragona, & Farber, 1993; Orlinsky & Howard, 1987). After and before the therapy hour, patients live their lives without the physical presence of the therapist. They spend much more time outside of psychotherapy sessions than within. It seems obvious that the experience of "being in therapy with the therapist" must be transferred into everyday life to yield sustained effects. Metaphorically, the ISP is the means by which patients "digest" what they experience in psychotherapy sessions.

The ISP is defined as "all conscious thoughts, memories, feelings, and fantasies related to therapy and the therapist that are experienced during the time *between* sessions." Orlinsky et al. (1993, p. 243) postulated that the ISP refers to patients' internalization of emotional, interpersonal, and cognitive experiences during psychotherapy sessions and that it connects the direct experience inside a session to the patient's experience and behavior outside of therapy, thereby linking it to outcome (Zeeck, Hartmann, & Orlinsky, 2004). Hence, it can be assumed that patients' psychological processing and active use of psychotherapy in the time between sessions will be influenced by (and in turn influence) the overall quality of the psychotherapeutic process (Orlinsky et al., 1993; Orlinsky, Rønnestad, & Willutzki, 2004).

Orlinsky and Geller (1993) described the following dimensions of the ISP, which can be measured by the Intersession Experience Questionnaire (IEQ; Orlinsky & Geller, 1993; Hartmann, Orlinsky, Geller, & Zeeck, 2003; Hartmann, Orlinsky, & Zeeck, 2011): (a) *intensity* of the ISP; (b) the *context* in which intersession experiences occur; (c) its *content* (comprising relationship fantasies, recreating the therapeutic dialogue, and applying therapeutic learning); and (d) its emotional tone.

According to previous investigations of eating-disordered patients (Hartmann, Orlinsky, Weber, Sandholz, & Zeeck, 2010; Zeeck & Hartmann, 2005), we expected that recreating the therapeutic dialogue and the emotional quality of the ISP (negative intersession emotions) would show significant relationships with outcome. ISP factors have been found to follow varying trajectories over time in therapy. For example, negative emotions (associated with thoughts of therapy or the therapist) are low at the beginning phase, relatively higher in the middle working phases, and lower again toward the end of therapies for patients not suffering from a personality disorder in inpatient treatments (Zeeck, Hartmann, & Orlinsky, 2006). In previous studies, ISP was repeatedly found to be related to outcome (e.g., Hartmann, Orlinsky, Weber, Sandholz, & Zeeck, 2010), but the designs and methods used thus far can be considered insufficient (e.g., no complete measurement over time; therapy samples of highly varying session numbers; aggregation of data instead of session by session modeling) to capture the expected complexity and nonlinear relations.

In the current study, we used the sample of a large multicenter randomized trial of outpatient psychotherapy for anorexia nervosa (the ANTOP study; Wild et al., 2009; Zipfel et al., 2014) to analyze the characteristics of the ISP and its relations to treatment outcome. The ANTOP study is the largest study to date on outpatient psychotherapy of adult patients with anorexia nervosa (AN). In the ANTOP study, cognitive behavioral and psychodynamic interventions

were compared with treatment as usual. To allow process–outcome analyses of the mechanisms of change and the conditions of failure or success, the ANTOP study included a comprehensive process measurement schedule covering up to 39 intersession intervals between the 40 regular treatment sessions. There was substantial weight gain in all conditions and in terms of weight gain, no differences between approaches were found. Overall, recovery rates after one year were relatively low (14%–19%; Zipfel et al., 2014). A better understanding of the treatment process and how these patients use psychotherapy could help to improve treatment strategies because psychotherapy is still the treatment of choice in AN (Herpertz et al., 2011; National Institute of Clinical Excellence, 2004).

Without insight into the complexity of ISP, it is difficult to explore in detail relations between process and outcome. The present study generated data with the potential to fill these gaps. Thus, our research questions were as follows: Do the facets of the ISP follow different trajectories over time? Are there significant differences in ISP between good or unfavorable outcome cases at 1-year follow-up? In pursuing these questions, we will present descriptive and graphical information about a comprehensive measurement of ISP in a randomized controlled trial and explore and test the relations of therapy outcome with the parameters of complex trajectory models.

Method

The ANTOP Study

The design, sample, and main outcome of the ANTOP study have been reported in detail elsewhere (Wild et al., 2009; Zipfel et al., 2014). Patients were randomly assigned to one of three outpatient treatment conditions: focal psychodynamic psychotherapy (FPT; $N = 80$); enhanced cognitive behavior therapy (CBT-E; $N = 80$); and optimized treatment as usual (TAU-O; patients received treatment by experienced outpatient psychotherapists and family doctors were asked to help with weight and physical risk monitoring; $N = 82$). Treatment in the FPT and CBT-E conditions comprised 40 sessions over a period of 10 months (starting with two sessions a week for the first two months, one session per week during months 3 to 6 followed by one session every other week until month 10). For a detailed description of the treatment manuals, see the overview in Zipfel et al. (2014) or the German manuals (Friederich, Herzog, Wild, & Zipfel, 2014; Legenbauer & Vocks, 2005).

This study involved 10 German university centers with experience in the treatment of eating disorders. Patient inclusion criteria were an age of ≥ 18, female sex, a Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV; American Psychiatric Association, 1994) diagnosis of anorexia nervosa or subsyndromal anorexia nervosa and a body mass index (BMI) of 15–18.5 kg/m^2. Exclusion criteria included comorbid conditions like substance abuse, psychotic or bipolar disorder, serious or unstable medical problems, the use of neuroleptic drugs, and ongoing psychotherapy. The main time points of measurement were at baseline (T0), after 4 months (T1), after 10 months (T2; end of FPT and CBT-E treatments), at month 13 (T3; 3-month follow-up), and at month 22 (T4; 12-month follow-up). For the following analyses, we used assessments at T0 and T4 to focus on overall symptom course and outcome. As primary outcome, we used the global outcome score defined in the original trial.

Measurement Schedule and Data Selection

All participants receiving FPT and CBT-E (manualized interventions) completed the IEQ before each psychotherapy session. Patients of the TAU condition were not included in the process measurement schedules. With the high frequency of measurement (every session, see below), it was possible to test if the process–outcome relations followed complex trajectories.

Participants

Of $N = 160$ individuals randomized for FPT and CBT-E, 153 completed the IEQ, and 108 (67.5%) could be analyzed because data on ISP as well as follow-up assessments at T4 were

Table 1
Sample Characteristics

	$N = 108$ N(%); $M(SD)$
Sociodemographic data	
Age	27.6 (7.9)
Marital status	
Single/never married	89 (83.9%)
Married over living as such	11 (10.4%)
Separated/divorced	4 (3.7%)
Unknown	2 (1.9%)
Clinical characteristics	
Initial BMI	16.7 (1.0)
Duration of illness	
< 6 years	69 (63.9%)
> 6 years	39 (36.1%)
Subtype	
Restrictive	56 (51.9%)
Binge–purge	52 (48.1%)
EDI total score	263.6 (50.1)
Remission class at T4	
Recovery	21 (19.4%)
Partially recovered	57 (52.8%)
Full syndrome	30 (27.8%)
Sessions attended	36.3 (7.6) Min = 9; Max = 41

Note. M = mean; SD = standard deviation; BMI = body mass index; EDI = Eating Disorder Inventory

available. There was no significant difference between patients with and without data on intersession experience with respect to age, marital status, total score on the Eating Disorder Inventory-2 (Paul & Thiel, 2005), duration of illness, subtype of anorexia nervosa, comorbidity, or BMI. For a sample description see Table 1.

Therapists

Treatment cases were conducted by 52 therapists ($N_{CBT} = 24$, $N_{FPT} = 28$), with a median of three cases per therapist and a range from 1 to 8. Because only 56% treated three or more cases, therapist effects could not be reliably estimated for this sample.

Measures

Trained assessors diagnosed comorbid conditions at intake (T0) and follow-up via the Structured Clinical Interview for DSM-IV (First & Gibbon, 2004; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 2004; Fydrich et al. 1997). Eating disorder-specific psychopathology was additionally assessed by a self-report measure, the Eating Disorder Inventory-2 (Paul & Thiel, 2005).

Therapy Outcome

BMI (kg/m^2). Height and weight were measured by independent study assistants at baseline and all main time points of measurement including T4. Participants were weighed in light clothing and without shoes.

The Structured Interview for Anorexic and Bulimic Disorders for DSM-IV and ICD-10 (SIAB-EX). The SIAB-EX (Fichter, Herpertz, Quadflieg, & Herpertz-Dahlmann,

1998) is a reliable and valid semistructured expert interview for the assessment and diagnostic categorization of eating disorders, conducted by trained assessors. The SIAB-EX was conducted at T0 and all further main time points of measurement (T1→T4).

Psychiatric Status Rating (PSR) Scale. With the PSR Scale (Herzog et al., 1993), masked assessors assessed overall severity of the eating disorder (T1→T4). Scores range from 0 (*no symptoms of AN*) up to 5 (*all criteria of anorexia nervosa fulfilled*) and 6 (*all criteria of anorexia nervosa fulfilled and severe impairment of functioning*).

Global outcome score/remission class. In line with the ANTOP study (Zipfel et al., 2014), three remission categories were distinguished to define outcome: recovery, partial syndrome, and full syndrome AN. PSR scores and BMI were combined to a global outcome score: Recovery was defined as a BMI > 17.5 kg/m² and a PSR score of 1 or 2, and full syndrome AN as having a PSR score of 5 or 6 and a BMI of ≤ 17.5 kg/m². All other individuals were classified as partial syndrome AN.

As our main endpoint, we choose the 1-year follow-up assessment because it should show whether treatment led to sustained effects and was internalized in a way that changed patients' thoughts and behavior in the long term. Additionally, pathological weight loss takes time and full relapses may not be completely apparent at the 3-month follow-up. Outcome results were defined and described in the primary publication of the trial (Zipfel et al., 2014).

Therapy Process

Therapy-related intersession process was assessed using the German version of the IEQ (Hartmann et al., 2003, 2011; Orlinsky & Geller, 1993). The IEQ contains 52 items, which participants completed before each psychotherapy session. Therapists were blind to patients' ratings. Previous research identified eight dimensions of patients' recollections, thoughts, and feelings about their therapist and therapy in the time between the last and the upcoming session.

The dimension of intensity (IEQ-A) reflects the frequency and duration of thinking about therapy and therapist. Two dimensions reflect the context in which thoughts and feelings about therapy arise: The emotive and problem-solving context (IEQ-B1), which includes times outside of therapy when one is feeling bad about something or thinking about how to deal with a problem, and the dreaming–drowsy states (IEQ-B2), which include doing routine activities, daydreaming, and times when falling asleep or waking up.

The content of patients' thoughts on these occasions occurs along three dimensions: recreating the therapeutic dialogue (IEQ-C1), which includes thinking over events of the last session and imagining a conversation with the therapist; relationship fantasies (IEQ-C3), which include patients thinking about their relationship with their therapist and daydreaming about the therapist; applying therapeutic learning (IEQ-C2), which includes trying to follow the therapist's suggestions and asking oneself what the therapist would advise when facing a difficult situation.

Finally, two dimensions reflect the emotions accompanying the patients' intersession experiences: positive emotions (IEQ-D1), which include feeling accepted, safe, grateful, relaxed, or confident; and negative emotions (IEQ-D2), which include feeling annoyed, embarrassed, discouraged, frustrated, or anxious. Items of the first subscale (IEQ-A) are rated on a 6-point scale ranging from 0 (*never think about it*) to 5 (*an hour or more*), whereas items of all other subscales are rated on a 5-point scale, ranging from 0 (*not at all*) to 4 (*very often*) in the German version of the instrument (Hartmann et al., 2003). The reliabilities of the factors (Hartmann et al., 2011) were deemed sufficient (all Cronbach's $\alpha > 0.7$, except for IEQ-B2, $\alpha = 0.53$). Table 2 lists descriptive statistics, Cronbach's α, and factor intercorrelations in this sample.

The ISP content factor relationship fantasies had little variation (factor score $\leq 1 =$ rarely in 83% of 4,722 session intervals) and was not included in subsequent analyses. Reliabilities for all but one factor were good. The factor dreaming–drowsy states included three items: (B1) "doing some routine activity," (B5) "falling asleep or waking up," and (B6) "in a dream." They seem to have good face validity, but their occurrence must not necessarily correlate.

Table 2
Psychometric Properties of the IEQ in this Sample

						Correlations						
IEQ section	IEQ factor	*M*	*Median*	*SD*	Cronbach's α	A	B1	B2	C1	C2	C3	D1
A Intensity	Intensity	2.36	2.29	0.70	0.77							
B Context	B1 Problem solving context	2.03	2.00	0.93	0.76	0.35						
	B2 Dreaming–drowsy context	1.30	1.33	0.83	0.62	0.43	**0.36**					
C Content	C1 Recreating dialogue	1.55	1.60	0.79	0.78	0.47	0.37	0.38				
	C2 Applying	1.74	2.00	0.93	0.80	0.12	0.33	0.13	**0.26**			
	C3 Relationship fantasies	0.57	0.40	0.65	0.80	0.23	0.24	0.22	**0.30**	**0.24**		
D Emotions	D1 Positive emotions	2.39	2.40	0.81	0.86	0.08	0.13	0.02	0.08	0.25	0.07	
	D2 Negative emotions	1.29	1.25	0.80	0.83	0.10	0.18	0.19	0.19	0.17	0.21	**−0.18**

Note. M = mean; SD = standard deviation; IEQ = Intersession Experience Questionnaire. Correlations of factor-z-scores, $z = (f - M_i/SD_i)$; where i = patient number). Boldface = within section correlations

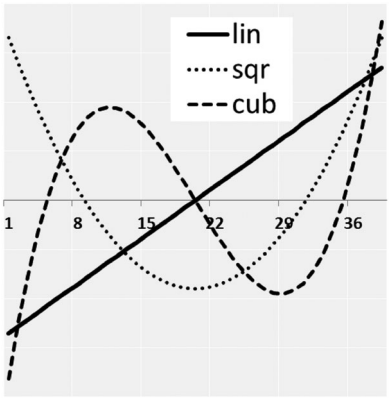


Figure 1a. Orthogonal polynomials.
Note. Y-Axis has no metric; X-Axis number of IS-interval; fits of individual trajectories are composites of the three orthogonal polynomials.

Data Analysis

From a methodologist perspective, the statistical analysis of ISP data is challenging because the process can follow multiple trajectories (including nonlinear or chaotic trends); it can have different lengths (number of sessions) or variably missing data; and, finally, the data are hierarchically structured (e.g., sessions and ISP variables within patients, outcomes on the patient level, and patients within treatment arms). In the present study, we decided to apply a variant of growth curve analysis (Bryk & Raudenbush, 1987; McArdle & Nesselroade, 2014), namely, growth curve analysis based on up to third-order orthogonal polynomials. Level-one trajectories for ISP variables were modeled with up to third-order orthogonal polynomials (Mirman, 2014) and fixed effects of level-two predictor variables. The models also included random effects of patients on all time terms. These and the fixed effects of the categorical outcome variable were added consecutively and their effects on model fit were evaluated by model comparisons.

Level-One models. For each patient an individual trajectory was fitted. The full model incorporates a linear, a quadratic, and a cubic component, all centered to the middle of the time axis. The centering of the components (to session 19 of 39) ensures that the resulting three dummy variables are uncorrelated ($|r| < 0.003$). Thus, the orthogonal polynomials constitute a set of statistically independent predictors avoiding problems of multicollinearity. Figure 1a shows examples of level-one fits as well as the overlay of the orthogonal polynomials.

The fitted trajectories have clearly distinct forms: linear, inverted U-shape, logarithm-like decelerating growth, or horizontal S-shapes. The full level-one regression formula comprises an intercept and three (first to third order) orthogonal polynomials (see Formula 1).

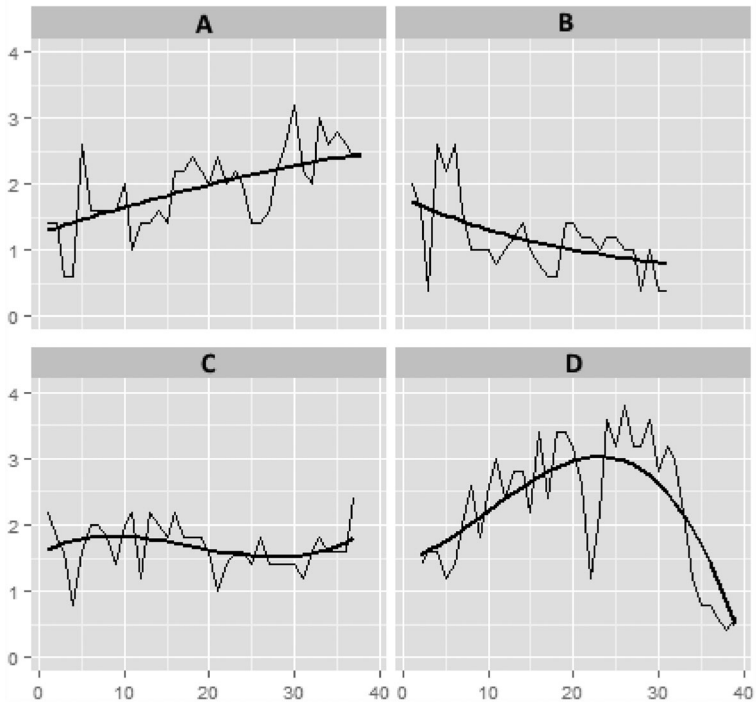


Figure 1b. Examples of raw scores and fitted curves.
Note. Y-axis = IEQ C1 recreating the therapeutic dialogue; X-axis = number of intersession interval (from 1 to 39). Panels A-F: raw scores and fitted trajectories of patients; A: example of predominantly linear trajectory; B: trajectory of a therapy finished before session 40; C: example of trajectory with cubic component (S-shape); D: example of predominantly quadratic trajectory (inverted U-shape).

Formula 1: Level-One Regression Model

$$Y = \beta_0 + \beta_1 \times Time + \beta_2 \times Time^2 + \beta_3 \times Time^3 + \varepsilon$$

We followed the suggestion of Mirman (2014, p. 47), that models can be restricted to third-order polynomials (maximum) as a compromise between sufficient curve fit and model complexity and found the fitting qualities of these three components very satisfactory (see Figure 1a).

Statistically, a mean group trajectory (intercept and slope, as fixed effect) as well as casewise deviations of its parameters (random effects) are estimated. We chose a statistical approach to determine model complexity: Starting with a linear base model, we added higher order polynomial terms to the model and tested whether they improve model fit, relying on the log likelihood ratio test, which tests if the addition of the model term significantly improves the preceding model (see Table 3, level-one model testing columns). The polynomials have no a priori psychological meaning and no meaningful scale. With every additional component, the fitted trajectories become more flexible, but the model ought to be as simple as possible and polynomials should be included only if they significantly improve the fit.

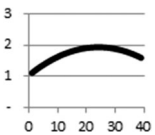
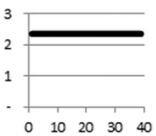
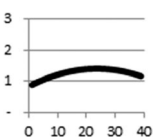
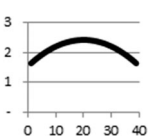
The first row in each ISP variable section represents the fit of a simple mean (intercept, no slope). All intercepts are significantly different from zero. The next three lines show the fit of trajectories when linear (first order), squared (second order), and cubic (third order) trends are added and their respective improvement statistics (chiSq-2 ΔLL). The respective p-value indicates if the fit significantly improved and if the additional component should be included or dropped from the model. The following columns show the estimates of all significant level-one curve parameters. To visualize the model, the resulting curve is shown in the iconic plot in the table.

Table 3
Level-One Models of ISP Factors

	Level-one model testing				Curve parameters[a]						
IEQ factor	Level-one model term	Chi-square −2 ΔLL	*df*	P <		Est.	*SE*	*df*	T	P <	Iconic plot[b]
A Intensity	Intercept				β_0	1.512	0.059	109.2	25.8	0.000	
	+ 1st order	110.2	4	0.000	β_1	−0.336	0.138	90.8	−2.4	0.017	
	+ 2nd order	54.5	5	0.000	β_2	−0.256	0.098	103.8	−2.6	0.010	
	+ 3rd order	1.4	2	0.489	β_3			Not estimated[a]			
B1 Problem-solving emotive context	Intercept				β_0	1.988	0.073	107.2	27.3	0.000	
	+ 1st order	123.3	4	0.000	β_1	0.180	0.137	98.1	1.3	0.193	
	+ 2nd order	70.6	5	0.000	β_2	-0.160	0.106	94.7	−1.5	0.134	
	+ 3rd order	0.7	2	0.718	β_3			Not estimated[a]			
B2 Dreaming–drowsy states	Intercept				β_0	1.246	0.066	109.5	18.8	0.000	
	+ 1st order	196.4	4	0.000	β_1	−0.357	0.141	89.9	−2.5	0.013	
	+ 2nd order	90.0	5	0.000	β_2	−0.151	0.104	92.4	−1.4	0.151	
	+ 3rd order	7.0	2	0.030	β_3	0.073	0.085	91.1	0.9	0.392	
C1 Recreating the therapeutic dialogue	Intercept				β_0	1.512	0.059	109.2	25.8	0.000	
	+ 1st order	167.2	4	0.000	β_1	−0.336	0.138	90.8	−2.4	0.017	
	+ 2nd order	64.1	5	0.000	β_2	−0.256	0.098	103.8	−2.6	0.010	
	+ 3rd order	3.2	2	0.200	β_3			Not estimated[a]			

(*Continued*)

Table 3
Continued

IEQ factor	Level-one model term	Level-one model testing: Chi-square −2 ΔLL	*df*	P <		Curve parameters[a]: Est.	*SE*	*df*	T	P <	Iconic plot[b]
C2 Applying therapeutic learning	Intercept			–	β_0	1.714	0.065	106.4	26.3	0.000	
	+ 1st order	445.1	4	0.000	β_1	0.894	0.179	91.2	5.0	0.000	
	+ 2nd order	218.0	5	0.000	β_2	−1.092	0.131	87.1	−8.4	0.000	
	+ 3rd order	0.9	2	0.630	β_3			Not estimated[a]			
D1 Positive emotions	Intercept				β_0	2.352	0.064	107.4	36.9	0.000	
	+ 1st order	309.0	4	0.000	β_1	0.221	0.131	87.0	1.7	0.096	
	+ 2nd order	144.0	5	0.000	β_2	0.107	0.102	87.9	1.1	0.294	
	+ 3rd order	0.4	2	0.801	β_3			Not estimated[a]			
D2 Negative emotions	Intercept				β_0	1.263	0.066	111.9	19.1	0.000	
	+ 1st order	603.1	4	0.000	β_1	0.516	0.140	94.9	3.7	0.000	
	+ 2nd order	116.6	5	0.000	β_2	−0.717	0.093	98.9	−7.7	0.000	
	+ 3rd order	0.1	2	0.939	β_3			Not estimated[a]			
Complex ISP negative recall	Intercept				β_0	2.147	0.169	112.8	12.7	0.000	
	+ 1st order	371.5	4	0.000	β_1	0.527	0.285	91.5	1.8	0.068	
	+ 2nd order	101.3	5	0.000	β_2	−1.537	0.228	96.6	−6.7	0.000	
	+ 3rd order	0.1	2	0.940	β_3			Not estimated[a]			

Note. df = degree of freedom; SE = standard error; BMI = body mass index; IEQ = Intersession Experience Questionnaire; ISP = intersession process. From various level one model fit statistics only the chi-square test of 2loglikelihood difference is listed due to limited space of presentation. Tables of complete statistical outputs are available from the authors.

[a]Parameters estimated only if corresponding polynomial is significant. See level-one model testing columns.

[b]Computed with significant level-one parameters only.

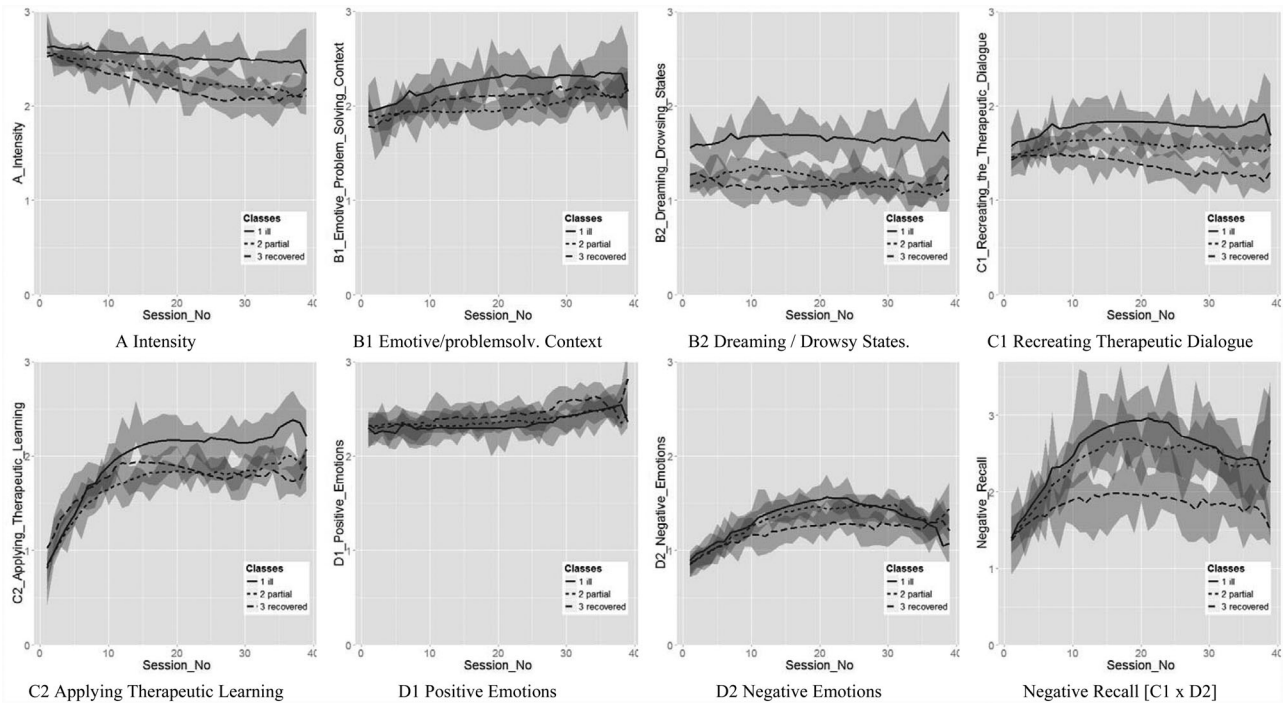


Figure 2. ISP trajectories by outcome category.
Note. Lines = mean of estimated curves by session and outcome; ribbons = ± standard error (SE) of raw data. The lines represent the means of the estimated curves and the SE ribbons represent the distribution of the raw; therefore, the curves and the SE ribbons can diverge in some time points. Negative recall = C1 recreating the therapeutic dialogue x D2 negative emotions. See Table 3 for difference testing. The axes of the intersession process factors are truncated to 3: theoretical maximum = 4; negative recall = C1 x D2; theoretical maximum = 16.

There are a few meaningful combinations of components. If there are no significant polynomials, the trajectory is represented by a horizontal line at the height of the intercept (if it differs significantly from zero). In that case, there is no process, but a stable trait-like state. A linear trend appears when the first-order polynomial is significant but second and third orders are not (intercept = session 19). A swinging S-curve appears when the third-order polynomial is significant. Depending on the size of the estimates, a combination of first- and second-order polynomials can appear as a log-like decelerating curve or have an inverted U-shape with an elevated end. Other combinations of significant estimates of the three polynomials are unpredictable because high degrees of freedom can fit many shapes. In any case, it is highly recommended to visually inspect the resulting fit.

Level-Two Models

Level-two models test if information given on the case level is predictive of the parameters of the individual curves. An individual curve (i) is defined by Formula 2.

Formula 2: Individual Trajectory

$$Y_i = \beta_{0i} + \beta_{1i} \times Time^1 + \beta_{2i} \times Time^2 + \beta_{3i} \times Time^3 + \epsilon_i$$

The classificatory predictor variable encodes the outcome at 12-month follow-up. It has three classes (recovery, partial syndrome, and full syndrome AN; further labeled "still ill"), in which the category still ill is treated as a baseline to contrast the other categories' parameters. Formula 3 represents all four possible predictions.

Formula 3(1–4): Level-Two Predictions of Level-One Intercepts and Slopes

$$(1)\ \beta_{0i} = \gamma_{00} + \gamma_{0Class} \times Class + \varepsilon_{0i}$$

$$(2)\ \beta_{1i} = \gamma_{10} + \gamma_{1Class} \times Class + \varepsilon_{1i}$$

$$(3)\ \beta_{2i} = \gamma_{20} + \gamma_{2Class} \times Class + \varepsilon_{2i}$$

$$(4)\ \beta_{0i} = \gamma_{30} + \gamma_{3Class} \times Class + \varepsilon_{3i}$$

Again, the model fit is tested for significant improvement for each additional level-two model term. The base model is the full level-one model, which includes first-, second-, and third-order time terms. We tested the relations of level-two predictors with all level-one terms. The analyses show if the outcome categories are significantly related to the betas of the first-, second-, or third-order terms (see Table 3).

The category of therapy failures (still ill) is contrasted against each of the other outcome categories. The parameters in these rows (still ill) comprise the base model, from which the two other outcome groups may differ significantly. A significant estimate for the category "Δ recovery" means that the slope of the fully recovered patients is significantly different from the slopes of the patients who were still ill at follow-up. The second- and third-order parameters were tested only if the respective predictors showed significant results in the previous steps of the analysis. All analyses were computed with R version 3.0.3 and the lme4-package (version 1.1-6). Parameter specific *p*-values are estimated with the lmerTest-package (Mirman, 2014, p. 57).

Results

The statistical method generates a potentially overwhelming amount of output and tests; therefore, (a) what is the most likely shape of a level-one trajectory? and (b) are there significantly different trajectories between groups?

Level-One Trajectories of the ISP Variables

To illustrate the diversity of the trajectories, all ISP factors are shown by outcome category in Figure 2. These mean trajectories clearly follow different composites of orthogonal polynomials. The factors intensity and positive emotions seem best represented by linear regression (first-order polynomial). Scores for dreaming–drowsy states show no significant trend, thus being best represented by constant values. The factors emotive and problem-solving context and applying therapeutic learning seem to follow decelerating log-like curves (approximated with composites of first- and second-order polynomials). Recreating the therapeutic dialogue resembles linear trend with slight log-like deviations, and, finally, the factor negative emotions shows inverted U-curves (second-order polynomial).

The incremental significance testing of model fit supports the visual impressions. With the very first session, the intensity of the ISP starts at a high level and slowly decreases. This process moderately decelerates approaching a medium level of intensity toward the end of therapy (see Table 2, intensity, intercept, first- and second-order slopes all significantly different from zero). The variables capturing the context of ISP showed very moderate trends, with barely noticeable but significant curvature. ISP in problem solving or emotive contexts shows a higher level compared to dreaming–drowsing states. It seems that these variables measure traits rather than states.

The content of ISP follows complex trajectories. Recreating the therapeutic dialogue slowly increases in the beginning, reaching a maximum in the first third of therapy and declines slowly toward the end. Applying therapeutic learning shows the steepest increase and approximates upper limits between the 10^{th} and 20^{th} session.

The emotional qualities of the ISP show very different trajectories. While positive emotions stay high and almost unchanged for the whole time of therapy, negative emotions clearly show an inverted u-shape. Furthermore, the model of positive emotions exhibited an unexpected behavior: Although adding a linear and a cubic term significantly improves model fit, the respective parameters are not significantly different from zero. Together with the exploratory graphical representation (Figure 2, D1), we conclude that it is likely to follow a flat line at high levels.

The results of the level-one modeling suggest that the ISP variables follow distinct and in some cases nonlinear trajectories over time. Intensity of the ISP is significantly but slowly decreasing over time, following linear trajectories. Both content variables of the ISP (recreating the therapeutic dialogue and applying therapeutic learning) are increasing and converging to a limit, but not like attaining a limiting value, rather like reaching an inflection point (most markedly for the process of applying therapeutic learning). Positive emotions show high levels from the very beginning with no significant change over time, whereas negative emotions showed an inverted U-shape.

Level-Two Outcome Categories and ISP

To identify the significant relations of the treatment outcome variable with the derived level-one curve parameters, level-two models were calculated as described above. See Table 4 for statistical tests and estimates.

Intensity (A). There are two significant differences between patients who are fully recovered at 1-year follow-up and patients who are still ill. Recovered patients start with a lower level of Intensity ($\Delta_{\text{intercept}} = -0.325$) and show a more negative linear slope ($\Delta_{\text{1st order slope}} = -0.761$). The visual impression of Figure 2 is supported by the significance testing: By the middle of time in therapy, successful patients had the least intense ISP, and partially recovered patients showed a similar but not significantly different trend.

Context (B1, B2). For the emotive and problem-solving context of the ISP, significant contrasts were found for the intercepts and slopes (still ill > recovery). For dreaming–drowsy states, there was a marked difference in overall level for the still ill patients, who showed the

Table 4
Level-Two Parameters

Variable	Parameter	Outcome category	Estimate	*SE*	*df*	*t*	*p*
A Intensity	Intercept	Still ill	2.508	0.118	110.83	21.207	**0.000**
		Δ Partial syndrome	−0.204	0.138	110.49	−1.481	0.142
		Δ Recovery	−0.327	0.155	108.71	−2.115	0.037
	Lin	Still ill	−0.452	0.253	93.82	−1.789	0.077
		Δ Partial syndrome	−0.442	0.294	93.31	−1.503	0.136
		Δ Recovery	−0.778	0.324	92.40	−2.404	**0.018**
	Sqr	Still ill	−0.098	0.187	91.34	−0.524	0.602
		Δ Partial syndrome	0.154	0.217	90.95	0.710	0.480
		Δ Recovery	0.279	0.239	89.44	1.167	0.247
	Cub	Still ill	−0.146	0.149	99.81	−0.980	0.329
		Δ Partial syndrome	0.126	0.174	99.44	0.722	0.472
		Δ Recovery	−0.001	0.193	97.74	−0.005	0.996
B1 Emotive problem-solving context	Intercept	Still ill	2.176	0.166	108.56	13.135	**0.000**
		Δ Partial syndrome	−0.256	0.194	108.27	−1.324	0.188
		Δ Recovery	−0.193	0.217	106.84	−0.891	0.375
	Lin	Still ill	0.572	0.311	101.08	1.838	0.069
		Δ Partial syndrome	−0.443	0.362	100.5	−1.225	0.224
		Δ Recovery	−0.528	0.398	99.38	−1.326	0.188
	Sqr	Still ill	−0.344	0.236	95.02	−1.456	0.149
		Δ Partial syndrome	0.431	0.275	94.62	1.570	0.120
		Δ Recovery	−0.191	0.302	93.16	−0.630	0.530
	Cub	Still ill	0.007	0.210	97.23	0.033	0.974
		Δ Partial syndrome	0.095	0.246	96.85	0.388	0.699
		Δ Recovery	−0.066	0.272	94.72	−0.244	0.808
B2 Dreaming–drowsy states	Intercept	Still ill	1.612	0.145	111.45	11.084	**0.000**
		Δ Partial syndrome	−0.146	0.328	91.66	−0.445	0.658
		Δ Recovery	−0.259	0.242	94.47	−1.070	0.287
	Lin	Still ill	0.018	0.188	90.85	0.097	0.923
		Δ Partial syndrome	−0.433	0.170	111.14	−2.548	**0.012**
		Δ Recovery	−0.494	0.190	109.48	−2.594	**0.011**
	Sqr	Still ill	−0.259	0.382	91.01	−0.678	0.499
		Δ Partial syndrome	−0.246	0.421	90.01	−0.584	0.561
		Δ Recovery	0.098	0.281	93.95	0.348	0.729
	Cub	Still ill	0.218	0.310	92.05	0.704	0.483
		Δ Partial syndrome	0.261	0.220	90.51	1.190	0.237
		Δ Recovery	−0.298	0.243	88.64	−1.223	0.224
C1 Recreating the therapeutic dialogue	Intercept	Still ill	1.745	0.129	112.10	13.524	**0.000**
		Δ Partial syndrome	0.183	0.291	97.38	0.627	0.532
		Δ Recovery	−0.252	0.240	97.22	−1.049	0.297
	Lin	Still ill	0.109	0.189	81.39	0.578	0.565
		Δ Partial syndrome	−0.207	0.151	111.71	−1.373	0.172
		Δ Recovery	−0.410	0.169	109.83	−2.431	**0.017**
	Sqr	Still ill	−0.323	0.339	96.73	−0.954	0.343
		Δ Partial syndrome	−0.971	0.373	95.34	−2.602	**0.011**
		Δ Recovery	0.027	0.279	96.75	0.097	0.923
	Cub	Still ill	0.107	0.307	95.22	0.348	0.729
		Δ Partial syndrome	0.001	0.221	81.09	0.007	0.995
		Δ Recovery	−0.026	0.244	79.71	−0.106	0.916
C2 Applying therapeutic learning	Intercept	Still ill	1.894	0.147	108.60	12.897	**0.000**
		Δ Partial syndrome	1.752	0.363	103.20	4.828	**0.000**
		Δ Recovery	−1.107	0.270	102.10	−4.102	**0.000**

(*Continued*)

Table 4
Continued

Variable	Parameter	Outcome category	Estimate	*SE*	*df*	*t*	*p*
	Lin	Still ill	0.612	0.265	85.59	2.307	**0.023**
		Δ Partial syndrome	−0.206	0.172	108.30	−1.201	0.232
		Δ Recovery	−0.184	0.192	106.50	−0.957	0.341
	Sqr	Still ill	−0.508	0.422	102.50	−1.203	0.232
		Δ Partial syndrome	−1.330	0.467	100.70	−2.850	**0.005**
		Δ Recovery	0.295	0.314	101.60	0.939	0.350
	Cub	Still ill	0.366	0.345	99.73	1.061	0.291
		Δ Partial syndrome	0.000	0.309	85.30	0.001	0.999
		Δ Recovery	0.027	0.344	83.33	0.079	0.937
D1 Positive emotions	Intercept	Still ill	2.240	0.145	109.10	15.451	**0.000**
		Δ Partial syndrome	0.231	0.287	95.77	0.804	0.423
		Δ Recovery	0.380	0.232	95.54	1.640	0.104
	Lin	Still ill	0.055	0.179	83.58	0.305	0.761
		Δ Partial syndrome	0.143	0.169	108.80	0.842	0.402
		Δ Recovery	0.152	0.190	107.40	0.798	0.427
	Sqr	Still ill	−0.055	0.334	95.09	−0.164	0.870
		Δ Partial syndrome	0.157	0.368	93.65	0.426	0.671
		Δ Recovery	−0.418	0.270	94.98	−1.552	0.124
	Cub	Still ill	−0.205	0.298	93.07	−0.688	0.493
		Δ Partial syndrome	0.000	0.209	83.30	0.001	0.999
		Δ Recovery	0.021	0.232	81.49	0.091	0.927
D2 Negative emotions	Intercept	Still ill	1.359	0.149	113.73	9.126	**0.000**
		Δ Partial syndrome	0.731	0.319	99.41	2.291	**0.024**
		Δ Recovery	−0.810	0.199	101.76	−4.059	**0.000**
	Lin	Still ill	0.156	0.148	97.07	1.054	0.294
		Δ Partial syndrome	−0.072	0.174	113.43	−0.415	0.679
		Δ Recovery	−0.188	0.195	111.90	−0.965	0.337
	Sqr	Still ill	−0.029	0.371	98.78	−0.077	0.939
		Δ Partial syndrome	−0.515	0.409	97.83	−1.258	0.212
		Δ Recovery	0.155	0.232	101.21	0.668	0.505
	Cub	Still ill	0.196	0.256	99.24	0.766	0.446
		Δ Partial syndrome	−0.061	0.173	96.71	−0.353	0.725
		Δ Recovery	−0.121	0.192	94.73	−0.631	0.530
Complex ISP negative recall	Intercept	Still ill	2.477	0.379	114.37	6.543	**0.000**
		Δ Partial syndrome	1.268	0.605	101.96	2.097	**0.038**
		Δ Recovery	−1.847	0.503	98.01	−3.669	**0.000**
	Lin	Still ill	0.381	0.433	91.73	0.881	0.380
		Δ Partial syndrome	−0.210	0.442	114.09	−0.474	0.636
		Δ Recovery	−0.717	0.496	112.73	−1.446	0.151
	Sqr	Still ill	−0.251	0.703	101.31	−0.358	0.721
		Δ Partial syndrome	−1.605	0.774	100.27	−2.073	**0.041**
		Δ Recovery	0.374	0.586	97.56	0.638	0.525
	Cub	Still ill	0.795	0.646	95.91	1.230	0.222
		Δ Partial syndrome	−0.074	0.505	91.43	−0.146	0.884
		Δ Recovery	−0.205	0.560	89.72	−0.367	0.715

Note. df = degree of freedom; SE = standard error; ISP = intersession process. Δ partial syndrome, Δ recovery = difference of parameter estimate compared to the still ill group.
[1]Not included because of insignificant level-one model test.
[2]Not included because of insignificant level-one slope estimate.
$p < 0.05$ in boldface.

highest frequency of this frame of recall (both intercept contrasts Δ partial and Δ recovery are significant) over the whole span of therapy.

Content (C1, C2). Compared to still ill patients, those who fully recovered recreated therapeutic dialogue less in terms of their level (intercepts): Their active recall diminished over time in therapy. The recovery group had a strongly reduced linear slope in applying therapeutic learning, compared to the still ill group. This leads to a lower plateau of the trajectories, a limit that occurs at about the 10th session. The still ill group showed a higher limit, and hit it only later, at about the 20th session (see Figure 2; C2).

Emotions (D1, D2). For both negative emotions and negative emotions, no significant differences between outcome categories could be found.

Negative recall. With regard to the composite score of negative recall, recovered patients showed a significantly different linear trend of the trajectory. The result of this is a flatter U-shape, compared to that of the still ill patients. The latter recall their therapies and therapists much more negatively in the middle phases of therapy time (see Figure 2).

To summarize, patients with good outcomes showed the *lowest* levels of applying therapeutic learning. Furthermore, no differences were found between good and bad outcomes concerning negative emotions or negative emotions in the ISP. Negative recall followed an inverted U-shape in all groups, although recovery patients showed the lowest levels.

Discussion

This study focused on patients' conscious thoughts, memories, feelings, and fantasies about therapy and their therapist between sessions (ISP) in a group of patients with AN receiving specialized outpatient psychotherapy. We explored how patients worked through psychotherapeutic treatments that were conducted in the context of a large randomized controlled trial. Treatments were specifically designed and manualized for AN, and therapist adherence was assessed at regular intervals. The controlled conditions and comprehensive assessments of patients can be considered a strength of the study.

Statistical Models for Trajectories

When examining characteristics of the ISP, results of growth curve modeling clearly showed that ISP facets follow different trajectories. Using orthogonal polynomials to fit trajectories is a useful strategy to fit curves with parsimonious sets of equations. Because there is no a priori clinical meaning in any of them (let alone their combinations), a visual inspection and interpretation of the mean trajectories of groups is required.

ISP

In terms of ISP intensity, patients' reports of thinking often about therapy and therapist decreased somewhat over time. This decrease might be related to a reduction in session frequency after 2 months and again after 8 months of therapy, or simply because of the fact that working with the same therapist over a longer period of time becomes less exciting.

Patients were more preoccupied with therapy contents in emotive and problem solving contexts compared to relaxed dreaming–drowsy contexts. This is something one would expect because psychotherapy generally encourages patients to use in-session experiences to deal with difficult situations outside of therapy and develop new ways of relating and coping (Owen, Quirk, Hilsenroth, & Rodolfa, 2012). The difference between levels of the ISP according to contexts was also found in a previous study (Zeeck et al., 2006). When considering the specific situation of patients with AN, one can assume that patients will define every meal and confrontation with food or their own body as a problem context in addition to many interpersonal situations. Thus, they are most likely to be confronted with multiple problem contexts during their waking hours.

The levels of applying therapeutic learning increased over time. It appears that patients need some time in the beginning of treatment until they start to apply experiences from therapy sessions in everyday life. This ISP facet comprises items that ask if patients try to solve problems in a way they had worked on them with their therapist–if they ask themselves *What would my therapist want me to do?* or if they did something the therapist had suggested. After 10 to 15 sessions, a certain level of applying therapeutic learning is reached that was held until the end of treatment. This might be a specific pattern for AN and/or dependent on the length of treatment: Owen et al. (2012) found that the number of sessions in brief treatments (6 to 10 sessions) was negatively associated with applying therapeutic learning in a mixed sample of students treated at a university counseling center.

In AN treatment and the ANTOP manuals (lasting 40 sessions), the first sessions of treatment have a strong focus on patients' ambivalence and motivation because a resistance to changing and a fear of engaging in close relationships are characteristics for this patient group. In this very first stage, therapists do not intend to promote change, in contrast to the brief interventions evaluated in the Owen et al. (2012) study. Additionally, applying therapeutic learning in the study of Owen et al. (2012) was related to the use of psychodynamic and interpersonal techniques, whereas CBT techniques showed no association. This possible relation between ISP facets and treatment orientation in the ANTOP study was not part of this investigation, but will be explored as a next step.

When testing for differences between outcome categories, we found that patients with a good outcome showed the *lowest* levels of applying therapeutic learning. Patients who are still ill 1 year after treatment kept trying to apply what they experienced in sessions more than patients in the other two outcome categories, especially during the second half of treatment. One reason could be that those patients were unable to effectively use the insights from therapy or suggestions of their therapist and so tried again and again, whereas the successful group succeeded and thus gradually reduced this facet of ISP. Other possible (and compatible) explanations could include that the therapist's interventions did not fit the patient's situation, or that patients remained too fearful or ambivalent to really try using their therapeutic experiences despite a persistent intention to do so.

For both applying therapeutic learning and intensity, one might expect that higher levels are related to good outcome, yet it seems that there may be an optimal corridor in which more is too much and/or indicative of an unproductive or otherwise problematic treatment process. A sufficiently good and productive ISP (e.g., of applying therapy) seems to involve a switch, in which more is not better but worse, perhaps as a sign of being stuck with no solution at hand. We see this switch occur in a certain phase of therapy. Just as good therapy outcome follows certain tracks (Lambert, 2010; Lutz, 2003), we may suppose that good process also follows certain tracks: In early therapy phases, it might be helpful to engage as well as explore new attitudes and action. Later in therapy, emerging success should allow one to use therapy continuously while engaging more in relationships and things happening in daily life that have nothing to do with anorexia nervosa and psychological problem areas.

In the last termination phase of therapy (e.g., the last five sessions), the plots of some trajectories show new turning points, broader confidence intervals, or other irregularities. This was very likely partly due to not-per-protocol termination and mostly due to characteristics of the termination processes (e.g., avoidance of persisting negative emotions, or anxiety and grief associated with separation). We had no theoretical model of the termination process and did not control for it statistically. This issue could be of importance for posttherapy stability of results and should be addressed in future analyses. These conjectures might be verified by further analysis of the in-session process data and/or qualitative analysis of session audiotapes.

Recreating the therapeutic dialogue occurred on a slightly lower level as compared to applying therapeutic learning, reflecting how much patients remembered events from the last session, thought about what they had been unable to express in the last session, or imagined a conversation with the therapist. This facet does not change much over the course of treatment, as if the dialogue with the therapist was permanently present, but not in a very conscious, intense form.

In terms of emotional quality of the ISP, positive emotion showed a constantly high level and with little variation. Patients in general seem to feel safe and supported by the treatment.

Negative emotions showed low levels and an inverted U-shape, with the highest level in the middle phase of treatment. This phase, according to treatment orientation, focuses on either behavior change or core conflicts and problem areas like attachment anxiety or self-esteem. It is the phase in which patients might be most challenged by treatment, feeling sometimes hurt, anxious, or frustrated. The inverted U-shape trajectory of negative emotions seems to be a general phenomenon because it was also found in other patient groups (Zeeck et al., 2006). Whereas patients with borderline personality disorder showed constantly high levels of negative ISP emotions throughout the whole treatment, patients without a personality disorder reported the highest levels in the middle phase of therapy.

Overall, our findings point to the fact that psychotherapeutic treatment can be considered a developmental process, with phases that differ in both therapeutic tasks and patient's process characteristics. In their study, Castonguay, Pincus, Agras, and Hines (1998) cited Stiles et al. (1990) and their assimilation model of psychotherapy to the effect that "things sometimes have to get worse before they get better" (p. 236). Patients have to face difficult and painful contents before understanding and exploring them, and before their mastery is possible.

Surprisingly, there were no significant differences in the level of ISP emotions when comparing outcome categories. Overall, AN patients seem to feel supported and safe with their therapists. They experienced negative ISP emotions at the same time, but this does not differ significantly between outcome categories. As positive ISP emotions were found to be correlated with a good therapeutic alliance (Hartmann et al., 2011; Owen et al., 2012), and good alliance with good outcome (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012), this common path must somehow be barred in this patient sample. However, like in a previous study on bulimia nervosa (Hartmann et al., 2010), recreating the therapeutic dialogue in combination with negative emotions (negative recall) was significantly more common in patients with a less successful course. This might indicate feelings of frustration and hopelessness in patients who are not able to use the therapeutic experience and the exchange with the therapist.

It is very likely that the reported results are typical for longer therapies only, with enough time for an arc of suspense (or a slope of boredom or rise of aversion). A different ISP will evolve in short or ultra short treatment and show other relations to outcome. There we might find stronger relations of alliance and positive intersession emotions with outcome and rarely persisting negative recall because of a stronger association with attrition. It is also very likely that other clinical populations will show a different ISP, which has already been shown for borderline personality disorder versus no personality disorder samples (Zeeck et al., 2006). Therefore, the results cannot be widely generalized without further study.

Limitations

Growth curve analysis with (up to third order) orthogonal polynomials was able to model very different individual trajectories. The visual inspection of all level-one fits (see Figure 1ab) showed excellent "splining" fit.

Testing three outcome categories as the only additional level-two covariate (fixed effect), we followed a straightforward statistical approach in which the statistical tests proved powerful enough to detect significant differences. Testing every ISP facet separately improved clarity and ease of interpretation, but that has its disadvantages such as multiple testing with potentially inflated significance levels. Given the exploratory nature of this investigation, our concern was not to miss process–outcome relations.

ISP was measured in a multivariate time series design, but growth curve analysis does not take into account lagged or direct interseries covariations or their relations to outcomes. Exploring the time series of both in-session and ISPs of the ANTOP process data is a matter for future analysis. ISP was measured retrospectively before each session. Therefore, processes of preparation for the session and expectations of the content of the session might have influenced or biased the recall. Methods of ecological momentary assessment (Wenzel, Kubiak, & Ebner-Priemer, 2016) should also be explored in future studies of ISP.

Conclusions

Overall, we were able to demonstrate a relationship between the intersession processing of therapy and outcome in outpatient treatment for AN, showing that the relationship very often is not linear and that different ISP facets follow different trajectories over time. Some ISP facets (e.g., ISP in dreaming–drowsy states) might be related to patient characteristics, whereas others (applying therapeutic learning, negative emotions) were dependent on treatment phase or were associated with more stable aspects of the therapeutic process like the therapeutic alliance (e.g., negative emotions). Exploration of differences between treatment arms of the ANTOP study and patient characteristics (e.g., initial BMI, chronicity, and subtype) seem promising topics for further analysis.

References

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Bohart, A. C., & Wade, A. G. (2013). The Client In Psychotherapy. In M. J. Lambert, Bergin and Garfield's handbook of psychotherapy and behavior change (6th ed.). Hoboken: John Wiley & Sons.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101(1), 147–158.

Castonguay, L. G., Pincus, A. L., Agras, W. S., & Hines, C. E. I. (1998). The role of emotion in group cognitive-behavioral therapy for binge eating disorder: When things have to feel worse before they get better. Psychotherapy Research, 8(2), 225–238. http://doi.org/10.1093/ptr/8.2.225

Fichter, M. M., Herpertz, S., Quadflieg, N., & Herpertz-Dahlmann, B. (1998). Structured Interview for Anorexic and Bulimic disorders for DSM-IV and ICD-10: Updated (third) revision. International Journal of Eating Disorders, 24(3), 227–249.

First, M. B., & Gibbon, M. (2004). The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II). In M. J. Hilsenroth & D. L. Segal (Eds.), Comprehensive handbook of psychological assessment, Vol.2: Personality assessment. (pp. 134–143). Hoboken, NJ, US: John Wiley & Sons Inc.

Flückiger, C., Del Re, A. C., Wampold, B. E., Symonds, D., & Horvath, A. O. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. Journal of Counseling Psychology, 59(1), 10–17. http://doi.org/10.1037/a0025749

Friederich, H.-C., Herzog, W., Wild, B., & Zipfel, S. (2014). Anorexia nervosa: Fokale psychodynamische Psychotherapie (1st ed.). Göttingen: Hogrefe Verlag.

First, M. B., Gibbon, M. G., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). Structured Clinical Interview for DSM-IV Axis II Personality disorders (SCID-II). Washington, DC: American Psychiatric Press.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). Structured Clinical Interview for DSM-IV-TR Axis I Disorders, research version, patient edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute.

Fydrich, T., Renneberg, B., Schmitz, B., Wittchen, H.-U., First, M. B., & Benjamin, L. (1997). SKID-II: strukturiertes klinisches Interview für DMS-IV; Achse II: Persönlichkeitsstörungen. [SCID-II: Structured Clinical Interview for DSM-IV; Axis II: Personality Disorders] Hogrefe, Verlag für Psychologie.

Hartmann, A., Orlinsky, D. E., Geller, J. D., & Zeeck, A. (2003). Der Inter-Session-Fragebogen (ISF) - Ein Instrument zur Erfassung von psychotherapierelevanten Prozessen zwischen Sitzungen. [Intersession Experience Questionnaire-an instrument to measure therapy related processes between sessions] Psychotherapie, Psychosomatik, Medizinische Psychologie, 53(11), 464–468.

Hartmann, A., Orlinsky, D. E., & Zeeck, A. (2011). The structure of intersession experience in psychotherapy and its relation to the therapeutic alliance. The Journal of Clinical Psychology, 67(10), 1–20.

Hartmann, A., Orlinsky, D., Weber, S., Sandholz, A., & Zeeck, A. (2010). Session and intersession experience related to treatment outcome in bulimia nervosa. Psychotherapy: Theory/Research/Practice/Training, 47(3), 355–370. http://doi.org/10.1037/a0021166

Herpertz, S., Hagenah, U., Vocks, S., von Wietersheim, J., Cuntz, U., & Zeeck, A. (2011). The diagnosis and treatment of eating disorders. Deutsches Ärzteblatt International, 108(40), 678–685. http://doi.org/10.3238/arztebl.2011.0678

Herzog, D. B., Sacks, N. R., Keller, M. B., Lavori, P. W., von Ranson, K. B., & Gray, H. M. (1993). Patterns and predictors of recovery in anorexia nervosa and bulimia nervosa. Journal of the American Academy of Child and Adolescent Psychiatry, 32(4), 835–842. http://doi.org/10.1097/00004583-199307000-00020

Herzog, W., Schellberg, D., & Deter, H. C. (1997). First recovery in anorexia nervosa patients in the long-term course: A discrete-time survival analysis. Journal of Consulting and Clinical Psychology, 65(1), 169–177.

Lambert, M. J. (2010). Using progress feedback to inform treatment: Conceptual issues and initial findings. In M. Lambert (Ed.), Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice (pp. 109–134). Washington, DC: American Psychological Association.

Legenbauer, T., & Vocks, S. (2005). Manual der kognitiven Verhaltenstherapie bei Anorexie und Bulimie [Manual of cognitive behavioral therapy for anorexia and bulimia] (1st ed.). Heidelberg: Springer.

Lutz, W. (2003). Efficacy, effectiveness, and expected treatment response in psychotherapy. Journal of Clinical Psychology, 59(7), 745–750.

McArdle, J. J., & Nesselroade, J. R. (2014). Longitudinal data analysis using structural equation models. Washington, DC: American Psychological Association.

Mirman, D. (2014). Growth curve analysis and visualization using R. Boca Raton: CRC Press.

National Institute of Clinical Excellence. (2004). Eating disorders. National Clinical Practice Guideline No. CG6. The British Psychological Society & the Royal College of Psychiatrists.https://www.nice.org.uk/guidance/conditions-and-diseases/mental-health-and-behavioural-conditions/eating-disorders

Orlinsky, D. E., & Geller, J. D. (1993). Patients' representations of their therapists and therapy: New measures. In Psychodynamic treatment research: A handbook for clinical practice (pp. 423–466). New York: Basic Books.

Orlinsky, D. E., Geller, J. D., Tarragona, M., & Farber, B. A. (1993). Patients' representations of psychotherapy: A new focus for psychodynamic research. Special section: Curative factors in dynamic psychotherapy. Journal of Consulting and Clinical Psychology, 61(4), 596–610.

Orlinsky, D. E., & Howard, K. I. (1987). A generic model of psychotherapy. Journal of Integrative and Eclectic Psychotherapy, 6(1), 6–27.

Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2004). Fifty years of psychotherapy process–Outcome research: Continuity and change. In M. J. Lambert & D. R. Dupper (Eds.), Bergin and Garfield's Handbook of Psychotherapy and Behavior Change (5th ed., pp. 307–390). Hoboken: Wiley John & Sons.

Owen, J., Quirk, K., Hilsenroth, M. J., & Rodolfa, E. (2012). Working through: In-session processes that promote between-session thoughts and activities. Journal of Counseling Psychology, 59(1), 161–167. http://doi.org/10.1037/a0023616

Paul, T., & Thiel, A. (2005). Eating disorder inventory - 2. Göttingen: Hogrefe.

Stiles, W. B., Elliott, R., Firth-Cozens, J. A., Llewelyn, S. P., Margison, F. R., Shapiro, D. A., & Hardy, G. (1990). Assimilation of problematic experiences by clients in psychotherapy. Psychotherapy, 27(3), 411–420.

Wenzel, M., Kubiak, T., & Ebner-Priemer, U. W. (2016). Ambulatory assessment as a means of longitudinal phenotypes characterization in psychiatric disorders. Neuroscience Research, 102, 13–21. doi:10.1016/j.neures.2014.10.018.

Wild, B., Friederich, H.-C., Gross, G., Teufel, M., Herzog, W., Giel, K., … Zipfel, S. (2009). The ANTOP study: Focal psychodynamic psychotherapy, cognitive-behavioural therapy, and treatment-as-usual in outpatients with anorexia nervosa–A randomized controlled trial. Trials, 23(10), 23.

Wittchen, H. U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (2004). SKID-I Strukturiertes Klinisches Interview für DSM IV - Achse I: Psychische Störungen. [SCID-I Structured Clinical Interview for DSM-IV-Axis I Disorders] Göttingen: Hogrefe.

Zeeck, A., & Hartmann, A. (2005). Relating therapeutic process to outcome: are there predictors for the short-term course in anorexic patients? European Eating Disorders Review, 13(4), 245–254.

Zeeck, A., Hartmann, A., & Orlinsky, D. E. (2004). Intersession Prozesse: ein vernachlässigtes Thema der Psychotherapieforschung. [Intersession Processes: A neglected topic of psychtherapy research.] Psychotherapie, Psychosomatik, Medizinische Psychologie, 54(6), 236–242.

Zeeck, A., Hartmann, A., & Orlinsky, D. E. (2006). Internalization of the therapeutic process: Differences between borderline and neurotic patients. Journal of Personality Disorders, 20(1), 22–41.

Zipfel, S., Wild, B., Groß, G., Friederich, H.-C., Teufel, M., Schellberg, D., … Herzog, W. (2014). Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): Randomised controlled trial. The Lancet, 383(9912), 127–137.